Assessing biodiversity:

Beyond the taxonomic impediment

Camila Duarte Ritter

2018

UNIVERSITY OF GOTHENBURG

Faculty of Science
Department of Biological and Environmental Sciences

Opponent: Peter Heintzman
Examiner: Bengt Oxelman
Supervisors: Alexandre Antonelli & R. Henrik Nilsson
© Camila Duarte Ritter

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without written permission.


Cover image: Photos from different places in Brazilian Amazonia. From left to right and up to down: terra-firme forest from a tower close to Manaus (credit: Aline Ramos); slam trap in igapós forest in Benjamin Constant; terra-firme forest in Juruá (credit: Yenneie Bredin); a *vitória-régia* in the Amazon river (credit: Mauricio Denardin); the meeting of Amazon and Negro rivers (credit: Renata Lemes); Jaú National Park on the Negro river; fungi in a terra-firme forest in Benjamin Constant; campina in Novo Airão; the sunrise in front of Cuieras river. Remaining photos by C. D. Ritter.

Copyright ISBN print: 978-91-7833-089-8
ISBN digital: 978-91-7833-090-4
Digital version available at http://hdl.handle.net/2077/54445
Printed by BrandFactory AB
Understanding nature is no longer enough, scientists have a moral duty to protect the subject of their study (unknown author).

This work is dedicated to Amazonia.
# Table of Contents

Abstract ...................................................................................................................................... 5
Svensk sammanfattning.............................................................................................................. 6
Publications included in this thesis ............................................................................................ 7
Additional publications not included in this thesis ................................................................. 8
Introduction .............................................................................................................................. 10
  Biodiversity assessment ....................................................................................................... 10
  Molecular tools ..................................................................................................................... 13
  Amazonian rainforest ............................................................................................................ 16
Objectives ................................................................................................................................. 21
Methods .................................................................................................................................... 24
  Sampling areas (Paper II, III, and V) .................................................................................... 24
  Sampling areas (Paper IV) .................................................................................................... 26
  Sampling strategy/Sample collection of soil and litter (Paper II, III, IV and V) ...................... 28
  Sampling strategy/Sample collection insects (Papers IV and V) ....................................... 28
  Physicochemical soil analyses ............................................................................................ 29
  DNA extraction of soil and litter. ......................................................................................... 31
  DNA extraction from the insect samples .............................................................................. 31
  PCR Amplification .............................................................................................................. 32
  Sequence analyses and taxonomic assessment .................................................................. 33
Results and discussion .............................................................................................................. 35
Conclusions and outlook .......................................................................................................... 39
Paper contributions ................................................................................................................... 41
References ................................................................................................................................ 42
Acknowledgements .................................................................................................................. 55
Papers ....................................................................................................................................... 58
  Paper I ................................................................................................................................. 59
  Paper II ............................................................................................................................... 68
  Paper III ............................................................................................................................... 82
  Paper IV ............................................................................................................................... 147
  Paper V ............................................................................................................................... 191
Abstract

Precise and accurate quantification of biological diversity is crucial for many fields of biological research and for understanding ecosystem services, biological interactions, biogeography and designing management strategies for conservation. The majority of current taxonomic knowledge is concentrated on a few groups of macro-organisms, mostly vertebrates and flowering plants, which represent only c. 0.7% and 3%, respectively, of the ~11 million estimated species of eukaryotes. The overwhelming majority of the extant biodiversity is challenging to sample and/or identify, which hinders biodiversity studies. While the world’s poorly known, inconspicuous organisms (e.g. fungi, insects, nematodes, and bacteria) are essential to understand the evolution, maintenance, and functioning of biodiversity, this taxonomic impediment is the strongest barrier to quantify the diversity of such groups. Recent molecular (DNA-based) tools, such as DNA metabarcoding, promise to speed up biodiversity quantification by several orders of magnitude at moderate costs. These methodological advances allow researchers to circumvent difficult, time-consuming specimen examination and identification, thereby accelerating biodiversity research in poorly studied groups and diverse ecosystems. In addition, these molecular methods make it possible to detect and identify rare and taxonomically challenging species as well as to quantify the biodiversity in virtually any location, which in turn could aid in conservation design and environmental impact assessment (EIA) studies. This thesis examines the potential and challenges of biodiversity assessments and to recommend methods to identify inconspicuous organisms. It also evaluates the use of such methods to assess the variation of biodiversity across space and ecosystems, and to identify the factors underling the uneven distribution of biological diversity in Amazonia. The results show that the major diversity patterns observed for macro-organisms in Amazonia do not hold true for all organisms. The results also highlight the complementarity of molecular and traditional taxonomic studies to better assess the biotic-abiotic factors that underpin the community composition and turnover of biological diversity.

Keywords: Amazonia; conservation; genetic diversity; metabarcoding; molecular sequences.
Svensk sammanfattning

Publications included in this thesis

The thesis is based on the following manuscript and papers, referred to in the text by their Roman numerals.


Additional publications not included in this thesis


Lentendu, G., Buosi, P. R. B., Cabral, A. F., Segóvia, B. T., Meira, B. R., Lansac-Tôha, F. M.,


Introduction

Biodiversity assessment

The term biodiversity is relatively recent, formally introduced in the "National Forum on BioDiversity", Washington, DC, 1986 (Wilson & Peter 1988). However, the need and desire to understand biological diversity is much older. A basic and widely used measure to quantify biodiversity is the number of species in a given area or species richness (Magurran 2004). There are currently 1.5 million eukaryote species described, and a recent estimate puts the number at 11 million extant eukaryotes on Earth, although other estimates range from 2 million to 1 trillion (Mora et al. 2011, Larsen et al. 2017). Impressive as these numbers are, they are dwarfed by the estimated prokaryote richness, estimated between 10^6-10^8 species (Amman et al. 1995, Locey & Lennon 2016, Schloss et al. 2016).

There is a strong bias in our understanding of biodiversity: due to the ease of sampling and identification in some taxonomic groups, nearly all of their expected diversity has been formally described. This is the case for birds were around 98% of the extant species are thought to be known (Bebber et al. 2007, Chapman 2009) and mammals, where around 99% of all extant species are thought to have been formally described (Chapman 2009). However, the overwhelming majority of the extant biodiversity does not belong to these groups. All vertebrates combined represent only 0.7% of the estimated number of eukaryote species, whereas 23-34% are represented by insects (Hamilton et al. 2010) and 14-28% by fungi (Fig 1; Hawksworth 2001, Taylor et al. 2014).
Figure 1. Comparison of the number of newly described species in Amazonia and the proportion of flowering plants and vertebrates in relation to all described eukaryotes. A) The number of new species described in Amazonia between 1999 and 2009 according to a WWF report (2013). B) Approximate proportion of the most well-studied groups of eukaryotes in relation to all of the Eukarya. The minute region in yellow represents all vertebrates combined (0.7% of all eukaryotes). The grey slice indicates the other eukaryotes (e.g. protists, nematodes and molluscs).

Species identification requires taxonomic expertise, which in turn represents a substantial prior investment of resources and time (Campbell et al. 2011). To catalogue all eukaryote species on Earth at the current speed of species discovery and description, some 1,200 years of additional research would be needed (Mora et al. 2011). In view of this, the decreasing costs of DNA sequencing combined with new methods for taxonomic assessment may offer an unique opportunity to characterize large sets of organisms or even entire communities using bulk samples or environmental DNA (Fig 2; Gibson et al. 2014). These advances came in the wake of DNA barcoding (molecular identification) of species and the existence of public and relatively highly populated reference sequence databases (Hebert et al. 2003). For some organisms, no prior taxonomic information is available (i.e. they are undescribed), or a complete taxonomic identification is impossible in the absence of sequence
data (as is often the case for bacteria and many fungi). Indeed, molecular data are often the only source of taxonomic information for many groups of microorganisms (Blaxter et al. 2005, Tedersoo et al. 2014, Vartoukian 2016).

**Figure 2.** Sampling strategy to sequence a large proportion of a site's full biological community using environmental DNA (eDNA) and bulk insect samples. (A) Study design and location of plots in one locality: at least three plots should be set up in each major vegetation type present (illustrated here by different shades of green) in the interest of statistical reproducibility; (B) Scheme for one plot, showing a Malaise trap (used to collect insect samples) in the middle, and twenty trees arbitrarily chosen for soil sampling within a 28-m radius (in red circles); (C) The Malaise trap will capture insects and the eaten plants and parasites in their bodies, while
the soil samples will provide environmental DNA (eDNA) for a large proportion of the habitat’s total biodiversity (e.g., soil organisms, roots, leaves, dead animals, and faeces). The soil sampling design is that of Tedersoo et al. (2014).

Molecular tools

Molecular data have revolutionized the study of inconspicuous organism groups. Nowadays, DNA sequence are an important source of information on micro-organismal evolution and ecology (Stajich et al. 2009, Bik et al. 2012, Loman et al. 2012, Creer et al. 2016). Biodiversity assessments using molecular methods form a powerful tool to understand entire communities and to understand their biotic and abiotic interactions. Even in highly diverse and poorly sampled environments for which reference databases are very sparsely populated, the use of molecular operational taxonomic units (OTUs; Blaxter et al. 2005) allows for assessment of genetic diversity and enables comparison among multiple sites (Stahlhut et al. 2013). DNA-based studies have the potential to overcome at least some taxonomic limitations and have been identified as a “transformative technology” for the entire field (Baird & Hajibabaei 2012). However, more effort and method development are needed to populate and correct errors in reference sequence databases and to understand the intrinsic differences between taxonomic and molecular biodiversity assessments.

There are many factors to consider when using molecular tools in biodiversity assessments, including DNA extraction, choice of genetic marker (e.g. Clarke et al. 2014, Elbrecht et al. 2016), sequencing method (e.g. Liu et al. 2013, Schirmer et al. 2015), and data analysis procedures (e.g. Beng et al. 2016). A serious caveat for using these molecular methods for biodiversity assessments is the lack of richly populated taxonomic reference databases, especially for the tropical regions of the globe. Without such reference databases, the recovered sequences cannot be assigned to resolved taxonomic levels. Furthermore, the molecular assessment is hampered by miss-annotated reference sequences, technically
compromised sequences (e.g., chimeras), and reference sequences annotated only at high
taxonomic levels (e.g. phylum; Kang et al. 2010, Nilsson et al. 2012). Some recent initiatives
try to mitigate these problems, including among others the Global Genomic Observatories
(Davies, Field & The Genomic Observatories Network 2012) and the UNITE database for
molecular identification of fungi (Abarenkov et al. 2016, Kõljalg et al. 2013). But even with
these efforts it remains challenging to correctly assess many newly generated sequences.

The lack of trustworthy reference sequences has implications for biodiversity
assessment. Beyond the obvious problems of erroneous characterization of biodiversity, the
use of high (unresolved) taxonomic levels (as a consequence of the partial nature of the
information available in reference databases) only provides limited biological information.
This is particularly problematic in the context of the current biodiversity crisis, with an
increasing loss of biodiversity at the species level (e.g. Scholes & Biggs 2005, Barbault 2011,
Giam et al. 2012), which has been identified as a major challenge to humanity in the next
decades (Steffen et al. 2015). Correct quantification and measurement of biodiversity are
prerequisites for its protection and for efficient conservation of natural ecosystems. However,
such quantification is not feasible for many taxonomic groups and most geographical regions
using traditional approaches. Considerable sampling and research efforts are necessary, but
these are often constrained by time, funding, logistical difficulties, and the availability of
taxonomic expertise (Campbell et al. 2011, Wheeler 2014).

The difficulty in assessing biodiversity is amplified in mega-diverse and poorly
sampled environments such as tropical rainforests (Balmford & Whitten 2003, Giam et al.
2012). For instance, hundreds of new species are described each year from the Amazonia
rainforest (Fig. 1; e.g. Del Hoyo et al. 2013, WWF 2013), and the rate of degradation of this
biome is alarmingly high (Fig 3; Soares-Filho et al. 2006, Malhi et al. 2008). This degradation
is probably driving many species to extinction before they are even discovered and formally
described (Wearn et al. 2012). In this way, even with the limitations of molecular methods, the use of unified protocols to describe diversity patterns and drivers are essential for fast and cost-efficient biodiversity studies (Paper IV).

Figure 3. Deforestation density in Brazilian Amazonia during 2015 - 2016. Map modified from Instituto Socio-Ambiental – ISA with data generated from: deforestation from INPE/PRODES (2016), conservation units and indigenous territories from ISA (2016), roads and Amazonian limits from IBGE (2012), and hydropower stations from ANEEL (2016). The data were collected between August 2015 and July 2016 and indicated almost 8,000 square kilometres of deforestation, an increase of about 30% in relation to the year before.
Amazonian rainforest

Amazonia is the largest tropical forest in the world. It comprises a tremendously high number of coexisting species, forming a very complex web of life linked together through biotic and abiotic interactions. At the same time Amazonia is considered fragile in the context of current human impact: relatively small alterations may have major impacts, such as the loss of ecosystem functionality (Malhi et al. 2008). Review and rigorous evaluation of current methods to protect biodiversity could therefore help to increase the efficiency of conservation.

One such method is the Environmental Impact Assessment (EIA), which is a technical and legal system for assisting in environmental management and for supporting sustainable development (Jay et al. 2007). Unfortunately, due to economic pressures, EIA requirements are often relaxed in some parts of the world, such as Brazil. It is therefore crucial to evaluate the performance and impact of systems such as the EIA, and outline new tools to improve the assessment and monitoring of diversity for the sustainable development of Amazonia and for the protection of its biodiversity (Paper I).

Quantifying biodiversity is crucial, but not enough. To understand how historical and ecological factors influence patterns of diversification and endemism, it is essential to identify the drivers generating and maintaining diversity. Several hypotheses have been suggested to explain the origin and distribution of the Amazonian biota (Leite & Rogers 2013). These hypotheses focus on processes that likely exerted different influences in different regions and periods of time (Ávila-Pires et al. 2007, Antonelli et al. 2010). Some examples include the zones of endemism between large rivers (Ribas et al. 2011); a west to east decrease in tree and animal diversity (Fig, 4; ter Steege et al. 2003, Pitman et al. 2001, Hoorn et al. 2010, Jenkins et al. 2015); and the diversity and evolution of vegetation types linked to particular soil characteristics, flooding regime, and nutrient availability. Biodiversity patterns may ultimately be linked to differences in geological formations and historical processes such as basin
formation (Hoorn et al. 2010), marine incursions (Webb 1995, Bates 2001, Lovejoy et al. 2006), and biological interactions (e.g. Chan et al. 2008).

However, all these patterns are primarily known from well-studied macro-organismal groups. In order to understand the distribution of all biodiversity and the factors that determine them, additional studies need to consider a much larger proportion of poorly known, inconspicuous but ubiquitous organism groups such as fungi, “protists” (various minute eukaryotes belonging to a range of different eukaryotic groups), and bacteria. These lineages are crucial to nutrient cycling, food-web dynamics, and host-pathogen processes (Stajich et al. 2009, Dominati et al. 2010, Mahé et al. 2017). The use of soil samples and universal genetic markers such as the 16S gene (prokaryotes) and the 18S gene (eukaryotes) would allow us to test general diversity patterns and their drivers (Fig 2; Paper II).
Amazonia comprises characteristic vegetation types (hereafter referred to as environmental types) that are closely linked to soil characteristics, flooding regime, and nutrient availability. Four widespread and important environmental types are terra-firme forests (Fig. 5A, 5B and 5C), várzeas (Fig. 5D, 5E and 5F), igapós (Fig. 5G, 5H and 5I), and naturally open areas (e.g. campinas; Fig. 5J, 5K and 5L). These environments support distinct sets of plants and animals, typically associated with mutually different kinds of soil.

Terra-firme forests are considered the most biodiverse environmental type and cover the largest area in Amazonia. They are unflooded and generally characterized by latosols (Falesi 1984). By contrast, várzeas and igapós are seasonally flooded forests. These areas cover approximately 5-7% of the Amazon basin (Peres 1997) and remain submerged during parts of the year, sometimes up to 240 days (Junk et al. 1989). Várzeas are flooded by white-water rivers that have their origin in the Andes, from where the floods transport large amounts of nutrient-rich sediments that are deposited in large floodplains. This makes the várzeas fertile areas (Junk et al. 2011). Igapós are flooded by black-water rivers that drain the pre-Cambrian Guiana shield, which is characterized by large areas of white sands (podzols). The water of these floods is transparent, with low quantities of suspended matter but with high amounts of humic acids, which give the water a brownish-reddish colour (Junk et al. 2011).

Finally, the open areas of Amazonia comprise non-forested areas dominated by grasses and shrubs (campinas), and low-canopy forests (campinaranas) (Anderson 1981), which together...
cover approximately 1.6% of the Amazon basin. Open areas resemble islands in a sea of forest and are related to nutrient-impoverished sandy soils (Prance 1996, Fine et al. 2005).

Figure 5. Major vegetation types in Amazonia. A - C: Terra-firme (non-flooded forest) is considered the most diverse environmental type in Amazonia. The photos show the high plant density and high canopy (around 40 m); D - F: Várzea forests (seasonally flooded forest) are considered the second most diverse habitat in Amazonia. The photos show the water colour and the high density of plants. F shows a water mark at ca. 15 m height in a várzea forest; G - I: Igapó (seasonally dry water streams). The photos show the low density of plants.
and high leaf accumulation in the soils; J - L: Campinas (white-sand areas). Typical features of campinas are sparse, small trees and exposed white sand soils, as can be construed from these photos.

Geology and soil physicochemical characteristics are often considered crucial for biotic dynamics, vegetation, and diversity patterns across Amazonia at local to regional scales (Vogel et al. 2009, Laurence et al. 2010, Higgins et al. 2011). For instance, diversity patterns and community composition have been associated with soil nutrients in plants (Laurence et al. 2010). Soil chemistry is associated with bacterial composition (Buckley & Schmidt 2001, Navarrete et al. 2013) and phosphorus, is associated with general microbial community composition (Faoro et al. 2010). For micro-organism diversity and composition, pH is a very influential factor (e.g. Osborne et al. 2011, Kuramae et al. 2012, Barnes et al., 2016). Even with several studies reporting the importance of soil characteristics on community structure, no unified pattern or scientific consensus has yet been achieved. Understanding the role of soil physicochemical compounds in shaping organism richness and composition of the world's largest rainforest is therefore crucial for identifying the factors underpinning biodiversity patterns (Paper III).

Molecular tools such as metabarcoding suffer from a set of limitations. For instance, metabarcoding methods have to compromise between taxonomic coverage and taxonomic resolution. When coupled with universal primers, 18S sequencing may capture the majority of eukaryote organisms. However, 18S is not variable enough to distinguish all eukaryotes at the species level (Hartmann et al. 2010, Lindahl et al. 2013). Additionally, for 18S the lack of reference sequences for the vast majority of extant eukaryotes is problematic (Guardiola et al. 2015). Furthermore, while is an excellent gene to target in environmental metabarcoding studies, no primer pair is truly “universal”. There is just too much nucleotide variability in potential primer sites (preventing primer binding; e.g Tedersoo et al. 2015). The use of other
genetic markers could increase the number of taxonomic groups detected or provide an improved taxonomic resolution for more specific groups. Such markers include the nuclear internal transcribed spacer region (ITS) for plants (Chen et al. 2010, Yao et al. 2010) and the cytochrome c oxidase subunit I (COI) for metazoans (Hebert et al. 2003). With better taxonomic resolution and coverage (e.g. use of multiple markers) it is possible to compare the diversity patterns obtained from environmental samples with those of well-studied groups such as birds and trees (Paper V).

**Objectives**

The aim of this thesis is to explore methods to assess biological diversity beyond taxonomic limitations in a diverse ecosystem with little reference data. Furthermore, I use the obtained data, covering a substantial part of the terrestrial biodiversity in the sampled regions, to describe diversity patterns and drivers in the world’s largest tropical forest. The thesis consists of two published and two submitted papers as well as one manuscript not yet submitted for publication. These studies have the following objectives:

**Paper I** – The goal of this study was to discuss the efficiency of Brazilian Environmental Impact Assessments in protecting the biodiversity of Brazilian Amazonia. To this end I assessed three of the most recent and largest infrastructure projects in Amazonia, with a particular focus on whether ecosystem threats and potential environmental impacts were properly assessed in accordance with the EIA principles. I also suggested three cost-effective complementary approaches (remote sensing, reflectance spectroscopy, and DNA metabarcoding) to complement the biodiversity assessment in EIAs.
**Paper II** – The aim of this manuscript was to test the patterns of biodiversity distribution in Amazonia, using operational taxonomic units (OTUs) derived from 39 soil and litter samples. I examined the data for correlation of eukaryote (18S) and prokaryote (16S) OTU richness and for OTU richness and composition patterns along a longitudinal gradient throughout Amazonia. I furthermore tested whether OTU richness and composition are correlated with Amazonian environmental types related to stress-level and nutrient availability.

**Paper III** – In this study, I aimed at comparing the relation between OTU richness, effective number of OTUs (abundance-based analysis), and community composition in different soils substrates: litter; the organic matter composed by animal debris, leaves, roots, and other organisms, and the mineral soil. Additionally, I aim to test whether OTU diversity (richness and abundance-based measures) and community turnover correlate with physical and chemical soil properties.

**Paper IV** – In this paper my goal was to test the potential use of bulk insect samples to assess the diversity of prokaryotes and eukaryotes that are associated with the sampled insects. I also aimed to assess to what extent soil diversity data are complementary to insect sample data in the context of producing more comprehensive biodiversity estimates of any particular site. I finally proposed a universal standardised protocol for metabarcoding studies - one that should be easy to apply in most terrestrial environments. To assess this, we tested our protocol in three countries (Brazil, South Africa, and Sweden) in different habitats.
**Paper V** – In this manuscript my goal was to compare the richness of environmental DNA data from insects, litter, and soil for three markers (16S – targeting prokaryotes; 18S targeting eukaryotes; and COI targeting metazoans) with the species richness of some of the most well-studied organisms in Amazonia: trees and birds, as assessed by traditional taxonomic studies. I analysed data from four localities in Amazonia in different environments (terra-firme, várzeas, igapós, and campinas). Additionally, I assessed whether similar patterns of diversity in Amazonia could be recovered with our richness data both from OTUs and from taxonomic based species richness.
Methods

Sampling areas (Paper II, III, and V) – Four localities were sampled in Brazilian Amazonia (Fig. 6). We chose these localities to maximize geographic distance and cover all major ecosystem types (terra firme, várzeas, igapós, and campinas). These localities are:

1 - Benjamin Constant – a municipality on the triple border of Brazil, Colombia, and Peru, situated approximately 1,100 km west of Manaus at the upper Solimões River (4°22′58″ S, 70°1′51″ W). This is a very difficult region to access (boat access only), with low population density and relatively low rates of deforestation. The region is situated in the south margin of the Amazonas river, comprising large areas of várzea forests, terra-firme forests, and some igapós forests from black water lakes;

2 - PARNA Jaú – this is a national park with an area of the 2,272,000 ha located in the lower Negro river (1°51′0.00″S, 61°36′60.00″W), 200 km northwest of Manaus, AM. There is a marked seasonality in the water level of Negro river. Generally, the low water season extends from September to November and the high water from May to August (Fundação Vitória Amazônica 1998). About 70% of the forest area is covered by terra-firme forest, which is characterized by large trees in the order of 25 m in height (Borges et al. 2001). There is considerable heterogeneity in local plant communities in the terra-firme forests, something that is related to soil mosaics in the region (Fundação Vitória Amazônica 1998). This heterogeneity may also in part be due to human disturbance (Ferreira & Prance 1999). About 12% is covered by igapó forests. We also collected samples from campinas in the municipality of Novo Airão, which is situated in the west margin of the Negro river and for the purpose of this thesis was considered the same locality;

3 - Reserva da Campina – situated 60 km north of Manaus, AM (2°35′30.26″S, 60°01′48.79″W). The reserve comprises approximately 900 ha, of which 6. 5 ha is stunted
heath forest (campinas) and tall heath forest (THF). The campina area (2.6 ha) is composed of a mosaic of shrub islands surrounded by white bare sandy soil. The canopy height is about 4-7 m for SHF with sparse trees (Luizão et al. 2007); and Reserva do Cuieras – which covers 22.7 hectares and is located about 70 km north of Manaus, AM (2°36’32.67”S, 60°12’33.48”W). The vegetation is a mosaic of evergreen forest with a canopy height of about 35 - 40 m, with emergent trees over 45 m tall, varying to open areas and igapós forest.

Analyses of the Igarape Asu catchment (Cuieiras Reserve area) indicated that valley forest environments (igapó and Campinarana) cover 43% of the area, whereas slope and plateau (terra-firme) forests occupy 26% and 31%, respectively (Zanchi et al. 2002). These two reserves are situated in the east part of Negro river and for the purpose of this thesis were considered the same locality;

4 - FLONA Caxiuanã – 371,000 hectares of rainforest, located 350 km west of Belém, PA (1°44’07”S, 51°27’47”W), in the lower Amazon region of northern Brazil. About 85% of the forest area is covered by terra-firme forest and about 10% by várzea and igapó forests. This reserve also has some campinas (Behling & Costa 2000).
Figure 6. Study area and sampling locations across Amazonia. Inset panels show details of each locality. A: Benjamin Constant; B: Jaú; C: Jaú, naturally open areas; D: Cuieras; E: Cuieras, naturally open areas; and F: Caxiuanã. The symbols in A–F represent different vegetation types characterised by different soil properties: circles = open areas; triangles = forest seasonally flooded by black water rivers; squares = unflooded forest; and crosses = forest seasonally flooded by white water rivers. The sampling strategy was designed to cover a wide longitudinal range in Amazonia. The map was constructed using QGIS (2012). Legend from Paper III.

Sampling areas (Paper IV) – In order to test the efficiency of all protocols, in paper IV we used a subsample of Brazilian samples, including all samples from Benjamin Constant (Fig. 7). Additionally, we sampled South Africa and Sweden as described below.

5 - South Africa – Seven plots were sampled with Malaise traps (Fig. 1F) kept open for 24 h. The plots were set up in south-western South Africa (33.51°S, 18.48°E), in a dry coastal area during August 2016. The Cape landscapes are dominated by the Cape Folded Belt (Deacon et al. 1992), and the Cape Floristic Region is considered a global biodiversity hotspot (Myers et al. 2000). The mean temperature in the sampled period ranges from 10°C to 18°C, and the mean precipitation is 24 mm/month (https://weatherspark.com);
**6 - Sweden** – Insects were collected with Malaise traps (Fig. 1E) in three forest fragments, three agricultural farms, and three pasture farms. The Malaise traps were deployed for a period of seven days in nine sampling localities. The sampling was done in June 2016. The samples were collected around the city of Skövde, an area characterized by relatively high biodiversity compared to other mainland regions in Sweden, and usually nutrient-rich soils derived from calcareous bedrock (58.44° N, 13.66° E). The temperature in the sampled period ranges from 10°C to 21°C, and the mean precipitation is 52 mm/month (https://weatherspark.com);

![Figure 7. Map of sampling localities (Paper IV).](image)

The sampling localities used for this study cover a wide geographical range comprising the Amazonian rainforest in Brazil, the western coast of South Africa, and central-south Sweden. In each locality, different environments were sampled to test the usefulness of our method.
in a wide range of habitats. Habitats are: TF = terra-firmes, VZ = várzeas, and IG = igapós in Brazil, PC1 = dry area, PC2 = costal area, and PC3 = costal area more urbanized in South Africa, and ARA = arable farms, FOR = forest fragments, and PAS = pasture farms in Sweden. The green gradient represents biomes described in Olson et al. (2001), ranging from densely forested areas (dark green) to open areas (light green). Legend from Paper IV.

**Sampling strategy/Sample collection of soil and litter (Paper II, III, IV and V)** – Three plots were sampled in each major vegetation type present in each locality (3–4 depending on the locality). We sampled soils by adapting the approach of Tedersoo et al. (2014) in order to minimize information loss while keeping comparability between the present and other studies. First, 20 trees were randomly selected within 28 m of the Malaise traps (Fig. 2). The collectors used gloves (changed between each sampled tree) and masks all the time to reduce the risk of sample contamination (Fig. 8A). We collected 40 litter and soil cores taken in two opposite directions from those trees and subsequently pooled soil and litter samples to obtain one soil and one litter sample for each plot. The litter consisted of all organic material above the mineral soil and varies from 0 to 50 cm of thickness (Fig. 8B). We collected litter with gloves that were changed between sampling of each tree. Thereafter, we collected soil in the same places, with the soil samples taken from the top 5 cm of the mineral soil using a metal probe with 2.5 cm of diameter (Fig. 8C). The probe was sterilized with fire after collecting soil from both sides of each tree to reduce the risk of cross-contamination among samples. The samples were stored in a plastic bag with the same weight of sterilized white silica gel, 1 – 4 mm thick, pre-treated for two minutes of microwave heating (800 W) and 15 min of UV light. All plots were tagged with GPS coordinates. All dry samples were processed at the University of Gothenburg, Sweden.

**Sampling strategy/Sample collection insects (Papers IV and V)** – Arthropods, mainly flying insects, were collected with Slam traps (Brazil) or Malaise traps (South Africa and Sweden). Both are tent-like traps made of fine mesh-netting, widely used in entomological studies aimed to capturing strong-flying insects (e.g. wasps, mosquitos, and
butterflies) that typically fly upwards after hitting a fine-scale net. Those insects are ultimately trapped in a bottle filled with ethanol at 96% concentration. These two types of traps differ mainly in shape, with Slam traps resembling an igloo (dome-shaped, Fig. 8D) and Malaise traps resembling a Canadian tent (higher on one end, Fig. 8E and F). We sampled for one day in Brazil and South Africa and for seven days in Sweden. The difference in sampling time was due to logistic limitations, and although this should affect the empirical diversity estimates, it should not compromise the evaluation of these methodologies.

Physicochemical soil analyses – We determined the physicochemical soil properties of each plot (from three soil subsamples totalling 117 samples). The pH was measured in water (soil:water ratio 1:2.5). The exchangeable concentrations were measured for sodium (Na), potassium (K), and phosphorus (P) using Mehlich-1 extraction (unit mg/dm³) and for calcium (Ca) and magnesium (Mg) using KCl (1 mol/L) extraction (unit cmolc/dm³). The sum of all exchangeable bases (SB, which comprises K⁺, Ca²⁺, Mg²⁺, and Na⁺; unit cmolc/dm³) was then calculated. We also estimated exchangeable aluminium (Al and H⁺Al; unit cmolc/dm³) extracted with calcium acetate (0.5 mol/L at pH 7.0), aluminium saturation index (m; unit %), and Base Saturation Index (V; unit %). The effective cation exchange capacity (t) as well as the cation exchange capacity (T) were measured at pH 7.0 (unit cmolc/dm³). The organic matter (M.O) was quantified (unit g/kg), and the C (organic carbon) was quantified using organic matter (M.O) = C (organic carbon) x 1,724 - Walkley-Black (unit g/kg). Soil texture was characterized as the percentage of fine (0.05 – 0.2 mm), coarse (0.2 – 2 mm), and total sand (0.05 – 2 mm) as well as the silt (0.002 – 0.05 mm) and clay (< 0.002 mm) fraction of the soil weight. We did not quantify nitrogen levels due to the highly volatile nature of nitrogen; its concentration changes quickly during sample storage due to the activity of soil bacteria and other microbes, and freezing the samples in our remote sampling localities was not feasible. All analyses were commissioned from EMBRAPA Ocidental (Brazil) and
followed the protocol described in Donagema et al. (2011). Afterwards, we used the mean of the three soil samples from the same plot to obtain a representative value for the measurement of each variable for each plot.

Figure 8. Photos from field sampling. A) Soil collection in a terra-firme in Bejamin Constant; B) Litter in an igapó in Cuieras Reserve. In this plot the litter depth was around 20 cm; C) Soil collection with the probe, we used the top five centimetres of soil; D) Slam trap used in Brazil; E) Malaise trap in a forest fragment in Sweden (photo credit: Karl Mauritsson); F) the Malaise trap used in the coastal area of South Africa (photo credit: Tobias Anderman).
DNA extraction of soil and litter – Ten grams (dry weight) of soil samples and 15 ml of the litter samples (corresponding to 3 - 10 grams of dry weight litter, depending on texture and composition of each samples) were processed for total DNA extraction using the PowerMax® Soil DNA Isolation Kit (MO BIO Laboratories), according to the manufacturer’s instructions. Each DNA sample was then concentrated and washed, following the manufacturer’s instructions. We added 0.2 ml of 5M NaCl and inverted 3-5 times to mix. We then added 10.4 ml of 100% cold ethanol and inverted 3-5 times to mix. We centrifuged at 2500 x g for 30 minutes at room temperature, decanted all liquid, and washed the DNA pellet with 70% cold ethanol. Residual ethanol was removed at ambient air temperature overnight, and the precipitated DNA was re-suspended in 1 ml sterile 10mM Tris buffer. The concentration of DNA was checked using a Nanodrop 2000c UV-Vis Spectrophotometer (Thermo Scientific) before and after the concentration and washing steps.

DNA extraction from the insect samples – We first tested the efficiency of five DNA extraction protocols on five insect samples obtained from Sweden (Paper IV). Since we found no significant differences among the protocols we decided to perform all subsequent analyses using a non-destructive protocol, namely the one described in Aljanabi and Martinez (1997), for insect sample extractions. The samples were immersed in 15 ml of salt buffer (0.4 M NaCl, 10 mM Tris-HCl pH = 8, and 2mM EDTA pH = 8) using a vortex mixer for 1 min. Then 1.5 ml of 20% SDS and 20 μl of 20mg/ml Proteinase K were added. Whenever the insects were not completely covered by the buffer, we added additional buffer with the same proportion of reagents until all insects were covered. The samples were incubated at 60°C overnight. After that, we transferred 15 ml of clear lysis solution into a new tube, and the insects were transferred into 99% ethanol for preservation. Then, 11.25 ml of 6 M NaCl was added to the lysis solution, and the samples were vortexed for 30 s at maximum speed. The samples were then centrifuged at 10,000 g for 30 min, and 25 ml of the supernatant were
transferred to a new tube. An equal volume of isopropanol was added. The samples were incubated at -20°C for 1h and then centrifuged at 10,000 g for 20 min at 4°C. Pellets were washed with 2 ml of 70% ethanol, dried in room temperature, and re-suspended in 300 μl of sterile dH2O.

**PCR Amplification – Ribosomal small subunit (SSU) 18S rRNA soil and litter:**

We targeted the V7 region of the 18S rRNA gene using the forward and reverse primers (5’-TTTGTCTGSTAATTSCG-3’) and (5’-TCACAGACCTTTATTGC-3’) designed by Guardiola *et al.* (2015) to yield 100-110 base long fragments. Amplification was performed in a total volume of 25 μl and consisted of: 0.25 μl of AmpliTaq Gold DNA polymerase, 5U/μl, 2.5 μl Pfu polymerase buffer 10x, 0.5 μl dNTP (final concentration of each dNTP 200 µmol; all above mentioned reagents are from Promega®, Sweden), 0.25 μl of 50 mol of forward and reverse primers, 20.25 μl of nuclease free water, and 1 μl of DNA template. The PCR started with an initial denaturation step of 2 min at 95°C and then 30 cycles of denaturation at 95°C for 1 min, hybridization at 50°C for 45 s, and elongation at 72°C for 1 min, followed by a final elongation at 72°C for 10 min and finishing at 4°C. Each sample was amplified three times and pooled to reduce biases of amplification efficiency variation on different species and stochastic effects of amplification (Carew *et al.* 2013, Edgar 2013, Piñol *et al.* 2015). The quality of the amplification was checked in UV light using GelRed™ stain (1%; Biotium, USA) on a 2% agarose gel. All samples were purified using the QIAquick® PCR purification kit. Dual PCR amplifications were performed for Illumina MiSeq sequencing (Illumina, USA), using fusion primers as described in Bourlat *et al.* (2016). For indexing, we used the Nextera XT DNA index kit (Illumina, USA) according to the manufacturer’s instructions. We checked the quality of the PCR products on a 2% agarose gel. We then made a size selection using magnetic beads and a magnetic stand, adopting the ratio 0.9:1 for beads/PCR product. We checked the DNA concentration in a Qubit 30® fluorimeter (Invitrogen, Sweden). We
assessed the quality and size selection of the PCR products with a 2200 Agilent 2200 TapeStation® (Agilent, USA). We normalized and pooled the PCR product (with the same concentration) following the Illumina protocol. The samples were sequenced at SciLifeLab (Stockholm, Sweden) using an Illumina MiSeq 2x250 machine. **18S insects:** We used the same primers as described above, however the amplification and sequencing was done by Macrogen (Republic of Korea) following standard protocols using the Illumina MiSeq 2x250 platform.

**Ribosomal small subunit (SSU) 16S rRNA:** We targeted the V3-V4 region (~460 bases) of the 16S rRNA gene using the forward primer (5’- CCTACGGGNGGCWGCAG - 3’) and reverse primer (5’- GACTACHVGGGTATCTAATCC - 3’) from Klindworth et al. (2013). **Cytochrome c oxidase subunit mitochondrial gene (COI):** We amplified a COI region of ~313 bases using an internal forward primer (5’ - GGWACWGGWTGAACWGTWTAYCCYCC - 3’, Leray et al. 2013) and the CO1 degenerate reverse primer (5’ – TAAACTTCAGGGTGACCAAARAAYCA - 3’, Folmer et al. 1994).

**Sequence analyses and taxonomic assessment** – We used the USEARCH/UPARSE v9. 0. 2132 Illumina paired reads pipeline (Edgar 2013) to filter sequence reads for quality, de-replicate and sort reads by abundance, infer OTUs, and remove singletons OTUs. We filtered the sequences to discard chimeras, and we clustered sequences into OTUs at a minimum similarity of 97% using a “greedy” algorithm that performs chimera filtering and OTU clustering simultaneously (Edgar 2013). We used SILVAngs 1. 3 (Quast et al. 2012) for assessment of the taxonomic composition of the OTUs, using a representative sequence from each OTU as query sequence. For the 16S and 18S data, we used the SINA v1. 2. 10 reference data for ARB SVN (revision 21008, Pruesse et al. 2012) for both markers. For COI we used
the sequences available in GenBank (Benson et al. 2005) as reference and blast with the “blastn” tool.
Results and discussion

In general, our results show that the spatial diversity patterns of macro-organisms observed in Amazonia are not the same for all organism groups. The incongruence between our environmental metabarcoding data and macro-organismal diversity patterns highlight the importance of complementary studies to arrive at general diversity conclusions. This is extremely important from the vantage point of conservation and to understand biotic-abiotic interactions that maintain ecosystem functionality. Below I discuss the main results of each paper of this thesis.

Paper I - My review of the Environmental Impact Assessment (EIA) of three large infrastructure projects showed that the current EIA system in Brazil is very general and imprecise in its requirements. Several reasons contribute to making biodiversity assessments based on EIAs inadequate, notably the limited spatial and temporal scopes. Due to the lack of sufficient biological characterization of multiple distinct environments, EIAs often fail to deliver a proper description of the species composition as well as a characterization of the abiotic environment, such as the quality of the soil and water. More should be done not only to assess and document biodiversity but also to provide suitable baselines for comparison between largely pristine and more altered environments. Similarly, more efficient mitigation strategies must be proposed. Many aspects should be considered to improve the quality of EIAs in Brazil: improving the training of researchers and practitioners; making the formal requirements of EIAs more specific and biologically sound; providing an improved definition of the terms of reference (the document that sets the minimal assessment necessary in each project); and requiring a more thorough inventory of the species in the areas expected to be directly and indirectly affected by the infrastructure projects. With these improvements, future studies are likely to become more analytical, less descriptive, and of enhanced usefulness.
**Paper II** - In this study I showed that the currently accepted diversity patterns in Amazonia do not hold true for diversity in general, suggesting a decoupling of biodiversity patterns between macro- and micro-organisms. It is urgent to characterize the diversity patterns also for micro-organism communities, which comprise the vast majority of the global biodiversity. After centuries of studies, our understanding of biodiversity patterns is still heavily biased towards a handful of charismatic organism groups. The assessment of tropical diversity is enhanced by the use of 16S and 18S markers, as explored alongside traditional inventories by taxonomists. However, many constraints and complications remain, notably the compromise between taxonomic coverage and taxonomic resolution, and the lack of well-populated reference sequence databases. But even with incomplete reference databases, the use of OTUs gives a reasonable community composition proxy. We found that the biodiversity patterns are characterized more strongly by habitats types than by the west to east gradient. The community composition is more similar between flooded forests (igapós and várzeas) than between other environmental types. The campinas and terra-firme also have a high number of shared OTUs, highlighting that the inundation stress could be a filter for biotic communities.

**Paper III** – In this study, I found that soil and litter richness, the effective number of OTUs (abundance-based analysis), and the community composition are related in prokaryotes but not in eukaryotes. These differences could be observed at the OTU level but not at the phylum and order levels. I also found that the dominant bacterial groups in our dataset differed from those reported as globally dominant in a recent study (Delgado-Baquerizo et al. 2018). Additionally, the physicochemical soil variables, notably organic carbon and pH, could predict, to a certain extent, soil and litter diversity across Amazonia. pH had a positive correlation with OTU diversity, which was expected given the generally low pH in Amazonian soils, varying between 3.65 to 5.14. I found a negative correlation between soil
organic carbon content and prokaryotic and eukaryotic OTU diversity. This result was unexpected since diversity is usually related to biomass in soils and litter. However, soil/litter diversity is also related to decomposition rates. With rapid decomposition, the bulk of the biomass may be bound to aboveground stocks. Soils are crucial for carbon cycling in terrestrial ecosystems, and our results suggest that a better understanding of the relationship between diversity (above and belowground) and carbon cycling may help modelling carbon deposition and biodiversity patterns. Furthermore, I found a significant effect of soil organic carbon content on community composition.

**Paper IV** – In this study I found a significant difference between the diversity of prokaryote and eukaryote organisms registered in insect and soil samples. This stresses the advantages of including insect samples to complement biodiversity assessments in soil metabarcoding studies. To achieve this goal we presented a combined protocol using Malaise traps and soil samples. I tested the protocol in a wide range of environments, from hot and humid rainforests, to species-rich Mediterranean meadows and natural and anthropogenic habitats in the temporal zone. All localities showed a strong difference in community composition between the two sorts of samples (insects and soils). I suggested the use of non-destructive DNA extraction in insect samples to preserve specimens for, e.g., subsequent taxonomic studies, which are fundamental to the discovery and description of new species. Adopting a massive and standardized sampling scheme would allow fast and cost-effective estimations of global biodiversity and complement traditional biodiversity inventories.

**Paper V** – In this study, I found no significant relationship between OTU and taxonomic richness. Furthermore, I did not recover the same spatial patterns of diversity between macro- and micro-organisms for any taxonomic group. This suggests that at small spatial scales, the biological characteristics of each taxonomic group and peculiarities of each environment may be more important than the general diversity patterns for explaining
differences in richness distribution. In particular, I observed higher-than-expected tree richness in terra-firme environmental types from the Cuieras locality. Furthermore, I found a higher-than-expected biological richness in metabarcoding data from campinas, which are assumed to have lower levels of macro-organismic diversity than the other environmental types (ter Steege & Hammond 2001, Fine et al. 2010, Borges & Almeida 2001, Draper et al. 2018). Previous studies already reported on the importance of campinas for beta-diversity (e.g. Draper et al. 2018), and my data suggest that these environments could be potential hotpots of microbial diversity.
**Conclusions and outlook**

The papers in this thesis identified some of the major challenges of biodiversity assessments and presented prospects and tools to describe biodiversity beyond the taxonomic impediment. The identified issues include difficulties in describing some organism groups; comparing biodiversity among localities with many undescribed species; and scarcely populated reference sequence databases. These studies focused on inconspicuous organisms, many of which are difficult to examine using traditional methods, and on the assessment and comparison of biodiversity in megadiverse, yet poorly studied, Amazonian ecosystems.

This thesis contributes to mapping Amazonian diversity in the context not only of charismatic macroscopic organisms, but of the majority of terrestrial organisms. With metabarcoding data, it was possible to test if the general diversity patterns for macroorganisms could be generalized to more extensive biodiversity data, as collected in this thesis. I hope that this thesis will advance the biodiversity assessment debate in terms of Amazonian diversity patterns and drivers. With the development and implementation of the molecular-based methods presented here, it was possible to quantify biodiversity without the need to spend decades to obtain a sizable batch of taxonomic information. This is of extreme importance considering the current biodiversity crisis and the socio-political instability in the countries that comprise Amazonia. We need to know and understand biodiversity to use our knowledge in areas such as ecology, conservation biology, landscape planning, and evolution.

Is important to highlight that, in general, molecular data derived from “environmental sequencing” should be seen as complementary to, rather than as competing with, traditional taxonomic studies. Indeed, a confluence of both lines of evidence is highly warranted, as it will be necessary to overcome their respective shortcomings. Currently, parts of the taxonomic community take little (positive) note of environmental sequencing studies and *vice
versa (Ryberg & Nilsson 2018). It is particularly important that results of environmental sampling are made available and communicated to taxonomists, because this could encourage work with lineages untreated in a taxonomic framework. This interaction is likely to create “taxonomy feedback loops” which would accelerate species and lineage discovery, and potentially add prominent branches to the tree of life (Nilsson et al. 2016, Tedersoo et al. 2017). Thereby such an integrative approach will also improve the efficiency of conservation strategies, as more effort could be directed to “true” comprehensive hot-spots of biodiversity, including taxonomic richness and unique lineages in groups “hidden” to approaches solely based on either taxonomy or environmental DNA. Conversely, taxonomists are in a position to increase the resolution in ecological studies by providing ample, richly annotated DNA sequence data of legacy species as well as all newly described species for use in molecular characterization of biological communities.
Paper contributions

**Paper I** – *Environmental impact assessment in Brazilian Amazonia: Challenges and prospects to assess biodiversity.* Camila Duarte Ritter (CDR) conceived of this study together with the co-authors and wrote the manuscript with contributions from all co-authors.

**Paper II** – *Locality or habitat? Exploring predictors of biodiversity in Amazonia.* CDR conceived of this study together with the co-authors, carried out all the field and lab work, performed all the analyses together with co-authors, and led the writing with contributions from all co-authors.

**Paper III** – *High-throughput metabarcoding reveals the effect of physicochemical soil properties on soil and litter biodiversity and community turnover across Amazonia.* CDR conceived of this study together with the co-authors, carried out all the field and lab work, performed all the analyses together with co-authors, and led the writing with contributions from all co-authors.

**Paper IV** – *Biodiversity assessment in the 21st century: An integrative protocol for assessing eukaryotic and prokaryotic diversity from soil and insect traps using high-throughput DNA metabarcoding in various environments.* CDR conceived of this study together with the co-authors, carried out all the field and lab work, performed all the analyses together with co-authors, and led the writing with contributions from all co-authors.

**Paper V** – *DNA metabarcoding shows distinct patterns of diversity for Amazonian macro- and micro organisms.* CDR conceived of this study together with the co-authors, carried out all the field and lab work, performed all the analyses together with co-authors, and led the writing with contributions from all co-authors.
References


Acknowledgements

After for years I could finish this thesis thanks to several people that are crucial during this time.

First of all, I want to thank my supervisor Alex Antonelli for the opportunity to develop this project, his exciting view of science and all good moments.

My supervisor Henrik R. Nilsson for all help with analysis, writing, English and to be a fountain of support all the time. I didn’t know if I would be able to keep my mental sanity without you.

Urban Olsson for all good moments, talks and possibility to live in a great place.

Daniele Silvestro to be a true friend, share many beers, cigarettes, funny moments and beyond have patience to read my texts and help me all the time I requested.

Soren Faurby for all funny moments and all help, even more in the last month. Thanks a lot!

Christine Bacon for all dinners, talks and be an amazing woman scientist that is an inspiration.

Bengt Oxelman, my examiner, for South Africa trip, discussions and help me to keep the project in track.

Alex Zizka, I don’t think I need to say how important you are, thanks to be part of my life and be one of the best friend I could have.

Toby Anderman, thank for all long talks, comprehension, patience and good shared moments, you are one the best person I meet.
Josué Anderson for all support, funny karaoke moments and to share this challenge of do a PhD out of our home.

Daniela de Abreu, my Mocambique sister, thanks a lot for everything; my life would be so harder and boring without you. Te amo neguinha.

Fabian Roger for all help and support, it was nice you that we could became friends (after the time that you didn’t like me…).

The latino guys, Pavel and Juan, to make my time more enjoyable.

Cintia Machado to share the house, wine nights, song duet, trips and so many good moments.

Ivan and Victor you made this place really funny and happy for me, thanks for be so great friends.

All Brazilian people that make this place more “home”: Maria do Céo, Climbiê, Beatriz, Romina, Thais, Thomas, Carine and Paola.

All my friends in Brazil for all support, talks and skypebeers: Linão, Lili, Gabi, Jonas, Ari, Juninho, Marcelão, Pancho, McCrate, Xu, Ju, Fe, Sassá and Guta.

Ylva Heed, for all support all the time, you are amazing.

Dom and Jonna for good talks and climbing time.

My Swedish friend, Yenn, for all moments, trips, makes me know more Norway than Sweden and to be present in my life for all this time, I am so happy to have you.

My family, Marta and Vini, to be my family.

My marine friend that I should hate bit you make it hard: Kristi, Elika, Emma and Olga.
All colleagues and formers to share so may moments and make it better: **Sven, Patrik, Harith, Yann, Allison, Ruud, Angela, Maria, Bernard, Claes, Johannes, Triranta, Sommah** and **Andrei**.

The work underlying this PhD thesis had primary financial support from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brazil: 249064/2013-8) and the Swedish Research Council (B0569601). I thank the Brazilian authorities ICMBio (registration number 48185-2) and IBAMA (registration number 127341) for the permits granted for this research; I would like to thank my colleagues and friends at the University of Gothenburg and the National Institute of Amazonian Research (INPA) for support and discussions, and field assistants, especially to **Emily Hagen** and **Yennie Bredin** for help in the collection phase. I also thank **Anna Ansebo** and **Sven Toresson** for all support; and all co-authors, for the productive and stimulating collaborations.
Papers